A fast and robust method for whole genome sequencing of the Aleutian Mink Disease Virus (AMDV) genome

Hagberg, Emma Elisabeth; Krarup, Anders; Fahnøe, Ulrik; Larsen, Lars Erik; Dam-Tuxen, Rebekka; Pedersen, Anders Gorm

Published in:
Journal of Virological Methods

Link to article, DOI:
10.1016/j.jviromet.2016.03.010

Publication date:
2016

Document Version
Publisher's PDF, also known as Version of record

Journal of Virological Methods 234 (2016) 43–51

Contents lists available at ScienceDirect

Journal of Virological Methods

journal homepage: www.elsevier.com/locate/jviromet

A fast and robust method for whole genome sequencing of the Aleutian Mink Disease Virus (AMDV) genome

Emma E. Hagberg[a,b,*], Anders Krarup[a], Ulrik Fahnøe[c,1], Lars E. Larsen[c], Rebekka Dam-Tuxen[a,2], Anders G. Pedersen[b]

[a] Kopenhagen Diagnostics, Kopenhagen Fur, Glostrup, Denmark
[b] Department of Systems biology, Technical University of Denmark, Lyngby, Denmark
[c] National Veterinary Institute, Technical University of Denmark, Frederiksberg, Denmark

Article history:
Received 9 October 2015
Received in revised form 23 March 2016
Accepted 23 March 2016
Available online 6 April 2016

Keywords:
AMDV
PCR
NGS
Whole genome sequencing

A B S T R A C T

Aleutian Mink Disease Virus (AMDV) is a frequently encountered pathogen associated with commercial mink breeding. AMDV infection leads to increased mortality and compromised animal health and welfare. Currently little is known about the molecular evolution of the virus, and the few existing studies have focused on limited regions of the viral genome.

This paper describes a robust, reliable, and fast protocol for amplification of the full AMDV genome using long-range PCR. The method was used to generate next generation sequencing data for the non-virulent cell-culture adapted AMDV-G strain as well as for the virulent AMDV-Utah strain. Comparisons at nucleotide- and amino acid level showed that, in agreement with existing literature, the highest variability between the two virus strains was found in the left open reading frame, which encodes the non-structural (NS1–3) genes. This paper also reports a number of differences that potentially can be linked to virulence and host range.

To the authors' knowledge, this is the first study to apply next generation sequencing on the entire AMDV genome. The results from the study will facilitate the development of new diagnostic tools and can form the basis for more detailed molecular epidemiological analyses of the virus.

1. Introduction

Aleutian Mink Disease (AMD), sometimes referred to as Plasmacytosis, is worldwide the most important disease in the mink farming industry. The disease affects mink of all ages and is caused by Aleutian Mink Disease Virus (AMDV), a single stranded DNA virus belonging to the family *Parvoviridae* (Bloom et al., 1980) genus *Amdoparvovirus* species *Carnivore amdoparvovirus* 1. Viral entry is respiratory, oral, or via the placenta (Broll and Alexandersen, 1996). Infection results in a harmful activation of the immune system leading to hypergammaglobulinaemia and systemic vascular diseases and glomerulonephritis. Animal welfare is reduced and infected animals either die due to organ failure or become persistently infected carriers transmitting the virus within and between herds (Decaro et al., 2012). Like other parvoviruses AMDV replicates only in dividing cells where it utilizes the host cell's transcription machinery. Multiple parvoviruses can infect the same host, and this is believed to contribute to the high recombination rate shown for parvoviruses compared to other DNA viruses (Shackelton et al., 2007). AMDV consists of two large open reading frames (ORF's); the left ORF (nucleotide 116–1975) coding for the non-structural (NS) proteins involved in gene regulation and replication, and the right ORF (nucleotide 2241–4346) coding for the viral capsid proteins (VP), and three smaller central ORF's (Alexandersen et al., 1988; Bloom et al., 1988). In Denmark AMDV is a pathogen that is monitored by a mandatory national control program (Danish Executive Order 1447 of 15/12/2009, 2009). Briefly, the program requires all farms to conduct screening of their animals at regular intervals according to the disease status of the region. Positive farms undergo a more intensive monitoring and are encouraged to depopulate followed by a thorough cleaning and disinfection. Given that parvoviruses are highly contagious and very resistant to environmental factors, managing AMDV imposes large costs on the fur

* Corresponding author.
E-mail address: eha@kopenhagenfur.com (E.E. Hagberg).
[1] Current address: Copenhagen Hepatitis C Program (CO-HEP), Department of Infectious Diseases and Clinical Research Centre, Hvidovre Hospital and Department of International Health, Immunology and Microbiology, Faculty of Health and Medical Sciences, University of Copenhagen, Denmark.
[2] Current address: Novo Nordisk A/S, Smørmosevej 17-19, DK-2880 Bagsværd, Denmark.

http://dx.doi.org/10.1016/j.jviromet.2016.03.010

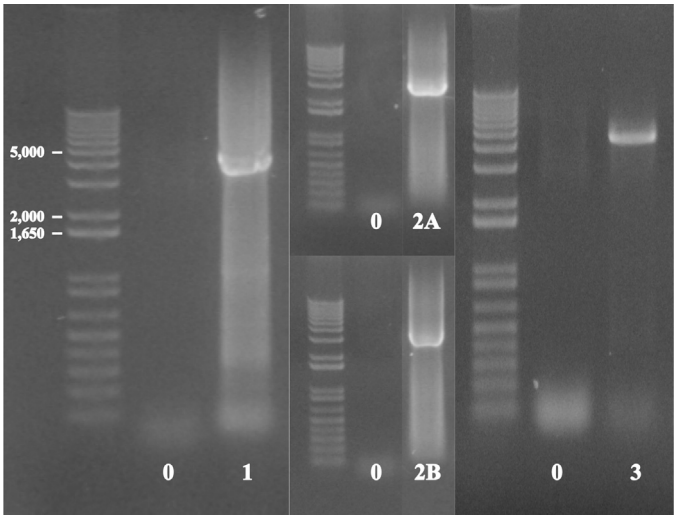

Fig. 1. Gel electrophoresis of PCR amplicons. A 1% agarose-gel showing the long-range PCR amplicons, fragment sizes are indicated with a 1 kb plus ladder. Lanes 0 represent negative control samples, lane 1 is AMDV-G amplified in one fragment (primers F1 + R3), lane 2A and 2B is AMDV-Utah amplified in two fragments; primers F1 + R2 and F2 + R3, respectively, and lane 3 is AMDV-G amplified in one fragment with primers F1 + R5.

industry (Decaro et al., 2012). The transmission patterns of AMDV between farms are not fully elucidated, and outbreak investigation is currently hampered by lack of sensitive tools for detection and typing of the virus. Previous studies have focused on smaller and conserved parts of the AMDV genome (Christensen et al., 2011; Knuuttila et al., 2015; Leimann et al., 2015; Oie et al., 1996) and therefore have produced data less suitable for typing. In addition, there are methods for characterisation of the AMDV genome using restriction fragmentation (Aasted, 1980) and Sanger sequencing (Alexandersen et al., 1988) but since they yield genetic information for limited stretches of the genome, they too provide less resolution than full genome sequencing. Next generation sequencing (NGS) is a powerful tool that has become cheaper and more easily available and it has successfully been applied to characterise entire genomes of other viruses and the genetic information obtained has been used to improve preventative measures (Escobar-Gutiérrez et al., 2012; Jakhesara et al., 2014; Kvisgaard et al., 2013).

To the authors' knowledge, the whole AMDV genome has not previously been sequenced using NGS. The aim of this study was to develop a fast, sensitive high-through-put method for full genome sequencing of the AMDV genome by NGS to lay the foundation for future development of tools for outbreak investigation, determination of virulence markers, and for development of more sensitive diagnostic tests and robust phylogenetic analyses.

2. Material and methods

2.1. Virus isolates

In order to establish an as universal method as possible two AMDV isolates with very different phenotypes, and hence presumably also different genotypes, were selected. The non-virulent cell-culture adapted strain AMDV-G (cell culture isolate, passage 10) was obtained from The Research Foundation of the Danish Fur Breeders' Association/Antigen Laboratory (Glostrup, DK), while the highly virulent AMDV-Utah isolate (antigen) was provided by emeritus Professor Bent Aasted (Copenhagen, DK). Total DNA was extracted using the QIAmp® MinElute Virus Spin Kit (Qiagen, Hilden, D) according to the manufacturer's instructions, and the final DNA elution was performed with 50 μL low TE-buffer.

Table 1
The primer sequences designed in the present study and the sizes of the expected amplicons for the applicable combinations. All primers have been designed with the Primer 3 software using the AMDV-G genome with accession number NC001662 as reference. Forward primers are indicated F, reverse primers R.

Primer	Positions (NC_001662)	Primer sequence (5′-3′)	Amplicon size (bp)
F1	77-97	CGCTTCGCGCTTGCTAACTTC	
R1	1814-1793	GCTCTGCGTGAGCGTTTGTTTC	
F1 + R1	77-1814		1735
F2	1502-1525	CCGGGGGGCACTGGAAAAACCTTG	
R2	3317-3296	GCAGAGAGGAGGTAGCCCCAAG	
F2 + R2	1502-2217		1816
F3	2934-2953	GCGTCGTTACAGGTTGCTTT	
R3	4467-4448	TTAATCCGCCACTTTCTGGT	
F3 + R3			1534
R5	4462-4439	CCGCCACTTTCTGGTAAAATAAGG	
F1 + R2			3240
F2 + R3			1946
F1 + R3			4390
F1 + R5			4385

2.2. DNA amplification

A long-range PCR covering about 91% of the AMDV genome (nucleotide position 98–4467, Table 1) was developed. The use of specific PCR amplification is important for future field applications as it avoids host DNA and attains a sufficient amount of double stranded DNA for the preparation of a sequencing library. The AMDV-genome was amplified in either a single or two overlapping fragments. The PCR primer-sequences are listed in Table 1 and were designed using the Primer3 software (Ye et al., 2012) with AMDV-G (accession no. NC001662) (Bloom et al., 1988) as reference genome. PCR reactions were setup as follows: 25 μL GoTag® Long PCR Master Mix (cat. no. M4021, Promega, Madison, WI), 2 μM primer F and R, 5 μL DNA template, and distilled water up to a total sample volume of 50 μL. Final cycling conditions were; initial denaturation at 95 °C for 2 min, 38 cycles of 30 s denaturation at 95 °C, 20 s annealing at 58 °C, and 30 s extension at 72 °C, followed by a final extension step for 10 min at 72 °C. All reactions were performed in a Bio-Rad CFX96 Touch instrument (Bio-Rad Laboratories, Inc., Hercules, CA).

The PCR products were analysed on 1% agarose gels stained with ethidiumbromide, and purified according to the manufacturer's instructions using the QIAquick® PCR Purification Kit or the QIAquick® Gel extraction kit (both from Qiagen, Hildren, D) depending on if there was a single product or not. DTU (Technical University of Denmark) Multi-Assay Core (Lyngby, Denmark) prepared the sequencing libraries according to the manufacturer's instructions and sequenced the samples on a 318-chip using the Ion Torrent PGM® (Life Technologies, Carlsbad, CA). The technology was chosen because it is easily accessible, cheap and fast, which are all-important factors for the future intended use in field applications. The region between positions 2470 and 2520 displayed very low read coverage in the Ion Torrent sequencing (possibly due to the nucleotide composition in this area) and we therefore Sanger-sequenced PCR-products spanning this region using primers F2 and R2 (Table 1) to verify the Ion-torrent findings.

2.3. Data analysis

Raw data in fastq-format were quality checked with FastQC version 0.10.1 and trimmed based on length (100–400 bp) and quality (average quality score >20) using Prinseq-lite (Schmieder and Edwards, 2011). Primer sequences were removed using Cutadapt version 1.4.1 (Martin, 2011). Reads were corrected for sequencing errors using RC454, and assembled with the associated Mosaik2

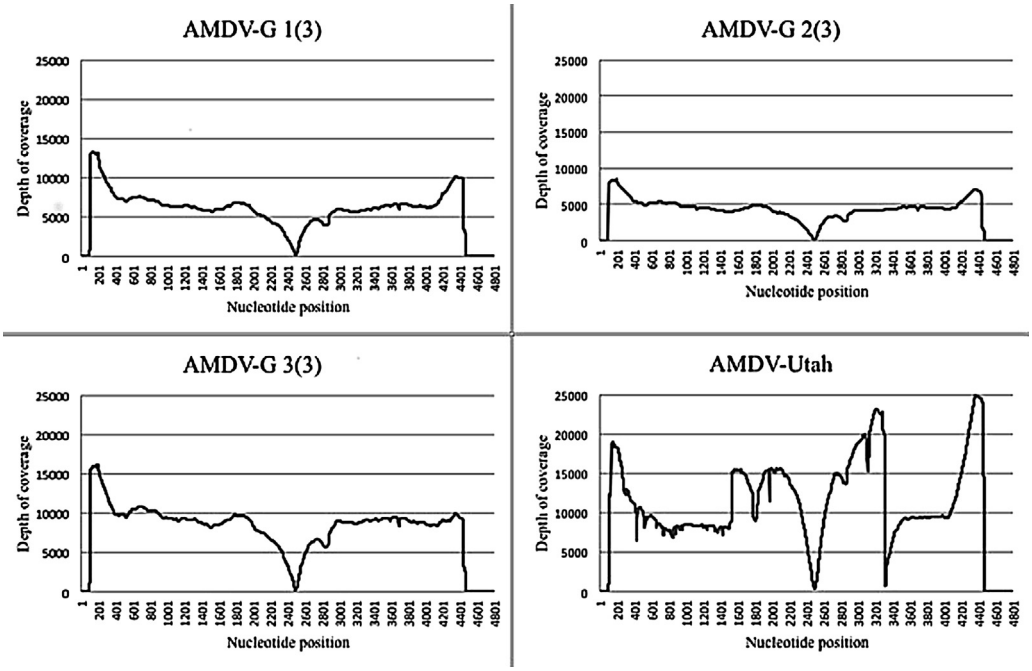

Fig. 2. Read coverage-plots. Coverage-plots showing the depth of coverage (*Y*-axis) at each sequenced nucleotide position in the genome (*X*-axis). The dip in coverage between position 2470 and 2520 is assumed to be caused by the ion-semiconductor sequencer having problems to read this homopolymeric region.

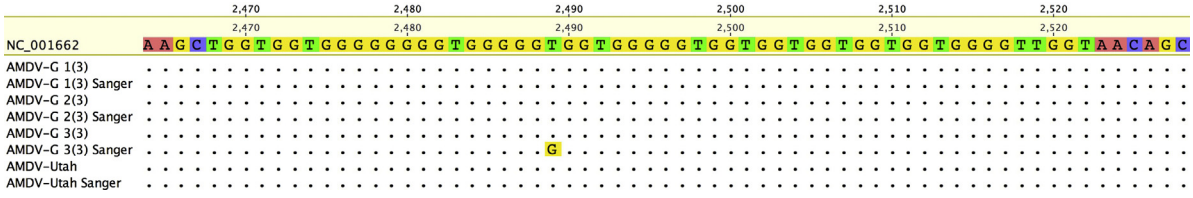

Fig. 3. Nucleotide alignment covering the homopolymeric region. Nucleotide alignment covering the homopolymeric region between positions 2470 and 2520 (repeated stretches of G's). Consensus sequences for each triplicate of AMDV-G, for AMDV-Utah, and for their corresponding Sanger generated sequence. The single nucleotide difference at position 2489 in the Sanger sequence for the third replicate of AMDV-G (3(3)) is highlighted.

assembler (Henn et al., 2012). The error-corrected reads were then mapped using BWA (Li, 2013) to the AMDV-G reference (NC001662) cut between the 5′ and 3′ annealing sites for primers F1 and R3 respectively (Table 1). For each sample a consensus sequence was generated using Vcf-tools 0.1.12a (Danecek et al., 2011). Multiple alignments of the full genomes was done using MAFFT v7.205 (Katoh and Standley, 2013). For additional comparisons, the following AMDV-Utah sequences were downloaded from the NCBI database: U39015.1, X77083.1 and Z1827.6.1. Alignments were visualized in Geneious 7.1.5. (Kearse et al., 2012).

3. Results

3.1. Specificity of the PCR products

A long range PCR assay for specific amplification of the AMDV genome was developed. Originally, three sets of primers spanning the AMDV genome were designed and the non-coding palindromic 5′- and 3′-ends (Bloom et al., 1990) which are known to interfere with PCR-amplification were excluded. After running a matrix with the possible combinations of forward and reverse primers, the optimal primer pairs was selected as assessed by gel electrophoresis. PCR cycling conditions were optimized by running an annealing-temperature gradient and amplification of PCR-products of the expected sizes were confirmed by gel electrophoresis (Fig. 1). The PCR reactions that produced one specific product were directly processed for sequencing. In cases where additional PCR-products were present, the band of the expected size was extracted from the gel prior to sequencing (Fig. 1). The AMDV-G and AMDV-Utah sequences sequenced in this study was amplified using primer F1 + R3 (Table 1), however during further assay optimisation better yields of PCR-product was achieved using primer F1 and R5 (assed by gel electrophoresis, Fig. 1). Despite the presumed genetic differences between the two viral isolates they both amplified well using these primers (Fig. 1).

3.2. Sequence quality and coverage

AMDV-G DNA was extracted in triplicates, and each individual sample was PCR-amplified and sequenced. AMDV-Utah DNA was extracted once, PCR-amplified, and sequenced. Primer-sequences and low quality reads were removed prior to data analysis. The data quality was overall high, and for each sample approximately 99% of the trimmed and quality filtered reads mapped to the AMDV-G reference. Coverage-plots for each of the four samples showed a dip in coverage between nucleotide positions 2470 and 2520 (Fig. 2). However, the Sanger-generated sequences spanning this region matched the sequences produced by the Ion Torrent, with the exception of a single nucleotide difference in the Sanger

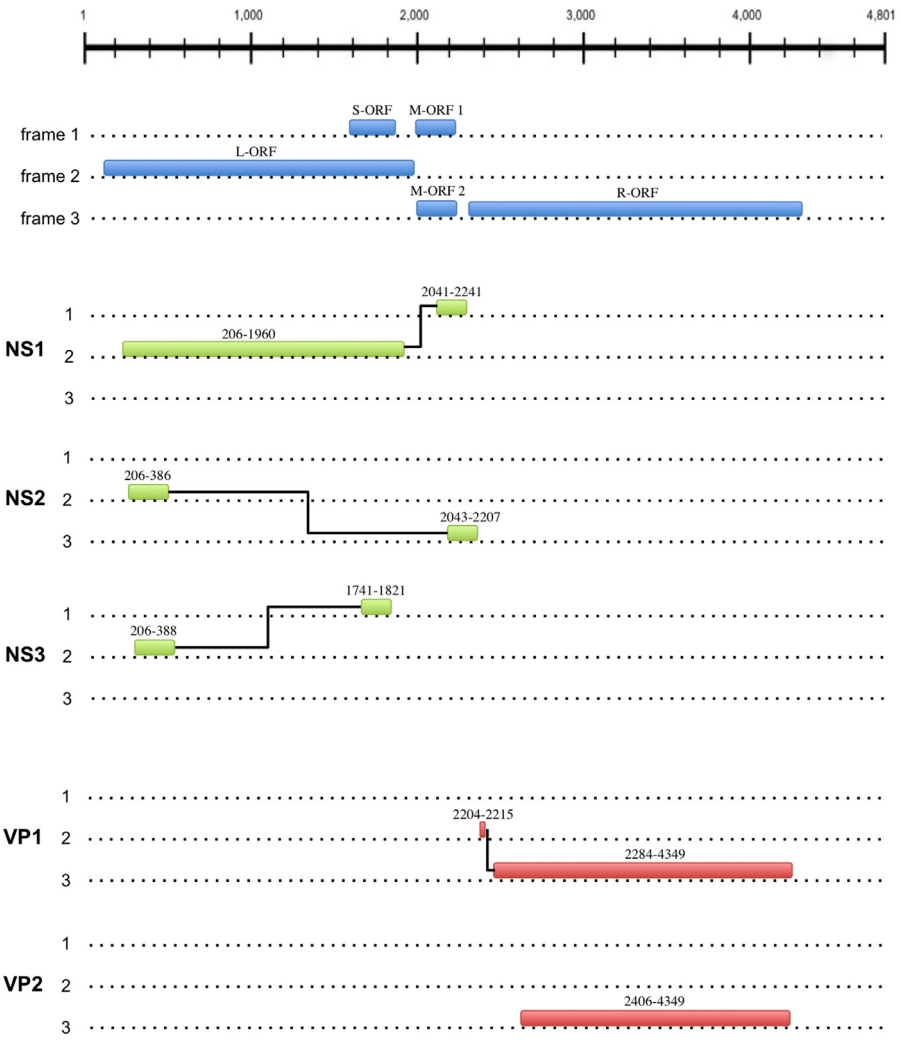

Fig. 4. Genomic map of AMDV. Genomic map of AMDV showing the open reading frames (ORF's) and how the major proteins are spliced together. Nucleotide positions refer to the reference AMDV-G genome. Nucleotide positions: 116–1975 left ORF (L-ORF), 1535–1825 small ORF (S-ORF), 1993–2209 mid ORF 1 (M-ORF 1), 1983–2204 mid ORF 2 (M-ORF 2), and 2241–4346 right ORF (R-ORF).

sequence for AMDV-G replicate no. 3 (Fig. 3). The quality score for this deviating base (G) was low (eight), while the scores for the corresponding base in the other two Sanger sequences were 38 and 40, and therefore indicates an error from the Sanger sequencing. The AMDV-G and AMDV-Utah sequences have the following accession numbers: KU513985, KU513986, KU513987, and KU513988, respectively.

3.3. Sequence analysis

The alignments showed an overall high sequence similarity (99% homology) between the AMDV-G samples generated in this study and the previously published AMDV-G sequence obtained by Sanger sequencing (NC001662). The robustness of the protocol was demonstrated by aligning the sequences obtained from sequencing AMDV-G in triplicates, and a deviation of a single nucleotide was found, out of the total 3369. An overview of the AMDV genomic organization is provided in Fig. 4. The AMDV-G triplicates and AMDV-Utah were compared on nucleotide and amino acid level (Table 2), and unless other is indicated, the results and discussion compare the AMDV-G and AMDV-Utah strains generated in the present study.

3.4. NS1 gene

The left open reading frame encodes for NS1–3. NS1 is the major regulatory protein in parvoviruses and it plays an important role in viral replication during infection (Fields et al., 2007; Gottschalck et al., 1994). In addition to confirming a number of previously reported differences between AMDV-G and AMDV-Utah, the present study also report novel findings as demonstrated in Table 2.

In the purine binding pocket between amino acids 421–492 (Gottschalck et al., 1994) a single change (F481L) between AMDV-G and AMDV-Utah was observed in addition to a single change (F430L) between ADMV-G and the AMDV-G reference (Table 2). Overall, the purine binding region, including the GKRN-region between amino acids 435–440 and its purine binding pocket, was well conserved in the sequences produced in the present study (Fig. 5, panel A). In agreement with previous studies of the distribution of changes in NS1 (Gottschalck et al., 1994), a higher degree of variability was demonstrated in the N- and C-terminals of NS1 compared to in the middle (Fig. 5 panel A).

Table 2
Overview of the nucleotide (nt) changes reported in the present study, including changes in the affected amino acids (aa) for the two most well described genes; NS1 and VP2. Absolute nucleotide positions refer to the AMDV-G reference genome (NC001662). Dash/es means no change in comparison to the AMDV-G reference genome. Isolates sequenced in the present study are indicated by *, and observations without a reference originates from the present study and hence are novel.

Absolute nt pos.	Codon pos.	NC001662		AMDV-Utah*		AMDV-G*		Reference accessions
		nt	aa	nt	aa	nt	aa	
L-ORF								
179–181		AGC	S	-	-	G--	G	
221–223	NS1/2 = 6	ATT	I	C--	L	---	-	Z18276.1, X77083.1
275–277	NS1/2 = 24	AAC	N	GCT	A	---	-	
290–292	NS1/2 = 29	GTT	V	C--	L	---	-	
293–295	NS1/2 = 30	GCC	A	--T	-	---	-	
299–301	NS1/2 = 32	TTG	L	C-A	-	---	-	
353–355	NS1/2 = 50	CCG	P	--A	-	---	-	Z18276.1, X77083.1
368–370	NS1/2 = 55	ACC	T	--T	-	---	-	
389–391	NS1 = 62	GCT	A	--A	-	---	-	
395–297	NS1 = 64	GAC	D	--T	-	---	-	
410–412	NS1 = 69	AAT	N	-CC	T	---	-	
416–418	NS1 = 71	ACA	T	-T-	I	---	-	
431–433	NS1 = 76	CAC	H	--G	Q	---	-	Z18276.1, X77083.1
440–442	NS1 = 79	AAC	N	--A	K	---	-	
443–445	NS1 = 80	AAT	N	G--	D	---	-	
470–472	NS1 = 89	TTG	L	--A	-	---	-	
485–487	NS1 = 94	CTG	L	G--	V	---	-	
491–493	NS1 = 96	ATT	I	G--	V	---	-	C--/L Z18276.1, X77083.1
506–508	NS1 = 101	AAA	K	--G	-	---	-	
509–511	NS1 = 102	AGC	S	--T	-	---	-	
524–526	NS1 = 107	AGT	S	GC-	A	---	-	-A-/N Z18276.1, X77083.1
527–529	NS1 = 108	AAC	N	G-T	D	---	-	
533–535	NS1 = 110	GTT	V	A--	I	---	-	
539–541	NS1 = 112	TTA	L	--C	F	---	-	
542–544	NS1 = 113	ATT	I	--C	-	---	-	
572–574	NS1 = 123	CAA	Q	--C	H	---	-	
650–652	NS1 = 149	TTT	F	--G	L	---	-	
653–655	NS1 = 150	ATG	M	--T	I	---	-	
659–661	NS1 = 152	AGA	R	-A-	K	---	-	
668–670	NS1 = 155	AAA	K	-G-	R	---	-	
680–682	NS1 = 159	GTT	V	-C-	A	---	-	C--/L Z18276.1, X77083.1
686–688	NS1 = 161	TAT	Y	-A-	F	---	-	
707–709	NS1 = 168	ATA	I	CA-	Q	---	-	
713–715	NS1 = 170	GAT	D	--C	-	---	-	
728–730	NS1 = 175	GAA	E	--G	-	---	-	
731–733	NS1 = 176	GAT	D	-CC	A	---	-	
734–736	NS1 = 177	AGA	R	-A-	K	---	-	
740–742	NS1 = 179	AAG	K	--T	N	---	-	
746–748	NS1 = 181	CTA	L	T-G	-	---	-	
767–769	NS1 = 188	GGA	G	--G	-	---	-	
776–778	NS1 = 191	AAG	K	--A	-	---	-	
788–790	NS1 = 195	TAT	Y	--C	-	---	-	
791–793	NS1 = 196	TTT	F	-A-	Y	---	-	
818–829	NS1 = 205	AAT	N	--C	-	---	-	Z18276.1, X77083.1
830–832	NS1 = 209	CAC	H	AC-	T	---	-	
836–838	NS1 = 211	AGA	R	--T	S	---	-	
845–847	NS1 = 214	ACA	T	GT-	V	---	-	
848–859	NS1 = 2015	TTC	F	A-A	I	---	-	
878–890	NS1 = 225	AAT	N	C--	H	-		
881–883	NS1 = 226	ACA	T	-AG	K	-	-	
884–886	NS1 = 227	GAT	D	--A	E	-	-	
887–889	NS1 = 228	AGT	S	G--	G	-	-	
902–904	NS1 = 233	TTT	F	-A-	Y	-	-	
920–922	NS1 = 239	GGC	G	--T	-	-	-	
923–925	NS1 = 240	ATT	I	--C	-	-	-	
925–928	NS1 = 241	GTT	V	A--	I	-	-	
953–955	NS1 = 250	AAA	K	--G	-	-	-	
954–056	NS1 = 251	ACT	T	G-C	A	-	-	
974–976	NS1 = 257	TTA	L	--G	-	-	-	
980–982	NS1 = 259	GAG	E	--A	-	-	-	
1007–1009	NS1 = 268	AAT	N	G--	D	-	-	
1025–1027	NS1 = 274	GGC	G	---	-	A--	S	
1070–1072	NS1 = 289	ACA	T	T--	S	-	-	
1106–1108	NS1 = 301	GCT	A	--A	-	---	-	
1109–1111	NS1 = 302	ACT	T	--C	-	---	-	
1130–1132	NS1 = 309	GAA	E	---	-	--C	D	
1154–1156	NS1 = 317	CCT	P	---	-	--G	-	
1181–1183	NS1 = 326	AGT	S	-A-	N	---	-	Z18276.1, X77083.1
1337–1339	NS1 = 378	ATT	I	--G	M	---	-	Z18276.1, X77083.1
1493–1495	NS1 = 430	TTC	F	---	-	C--	L	
1646–1648	NS1 = 481	TTT	F	C--	L	---	-	Z18276.1, X77083.1
1703–1705	NS1 = 500	GAC	D	--T	-	---	-	Z18276.1, X77083.1

Table 2 (*Continued*)

Absolute nt pos.	Codon pos.	NC001662		AMDV-Utah*		AMDV-G*		Reference accessions
		nt	aa	nt	aa	nt	aa	
2143–2145	NS1 = 620	TGC	C	---	-	-A-	Y	
2161–2163	NS1 = 626	AGT	S	G--	G	---	-	Z18276.1, X77083.1
M-ORF1								
2143–2145	NS1 = 620	TGC	C	---	-	-A-	Y	
2161–2163	NS1 = 626	AGT	S	G--	G	-	-	Z18276.1, X77083.1
								2179: GCA/A to CCG/P Z18276.1, X77083.1
2218–2220	non-coding	ATA	I	G--	V	G--	V	Z18276.1, X77083.1
M-ORF2								
2142–2144	NS2 = 94	CTG	L	---	-	--A	-	
2160–2162	NS2 = 100	GAG	E	-G-	G	---	-	Z18276.1, X77083.1
2217–2219	non-coding	AAT	N	-G-	S	-G-	S	Z18276.1, X77083.1
R-ORF								
2631–2633	VP2 = 76	GAC	D	--T	-	---	-	Z18276.1, U39015.1
2673–2675	VP2 = 90	AAA	K	---	-	C--	Q	
2679–2681	VP2 = 92	CAT	H	---	-	---	H	GC-/A Z18276.1, U39015.1
2685–2687	VP2 = 94	CAA	Q	---	-	---	-	A--/K Z18276.1, U39015.1
2748–2750	VP2 = 115	TAT	Y	-T-	F	---	-	Z18276.1, U39015.1
2751–2753	VP2 = 116	ATA	I	T--	L	---	-	
3459–3461	VP2 = 352	ATT	I	G--	V	---	-	Z18276.1, U39015.1
3585–3587	VP2 = 394	CAA	Q	---	-	--G	-	
3588–3590	VP2 = 395	CAC	H	--G	Q	A--	N	Z18276.1, U39015.1
3693–3695	VP2 = 430	TAC	Y	---	-	--T	-	
3696–3698	VP2 = 431	TAC	Y	---	-	ATT	I	
3705–3707	VP2 = 434	AAT	N	CAT	H	---	-	Z18276.1, U39015.1
3876–3878	VP2 = 491	AAC	N	G--	D	---	-	GAG/E Z18276.1, U39015.1
3975–3977	VP2 = 524	CCG	P	--A	-	--A	-	Z18276.1, U39015.1
4005–4007	VP2 = 534	CAT	H	G--	D	---	-	Z18276.1, U39015.1
4125–4127	VP2 = 574	AAT	N	---	-	G--	D	
4263–4265	VP2 = 620	AAG	K	AAC	N	---	-	Z18276.1, U39015.1
4305–4307	VP2 = 634	ATA	I	---	-	--G	M	

3.5. VP2 gene

In addition to confirming a number of previously reported differences between AMDV-G and AMDV-Utah, some of which have been proposed to influence virulence and host range (overview in Table 2), this study report novel differences in the VP2 gene. The N-terminus of VP2, amino acid 1–220, has been suggested to play a role in AMDV host range and culturing ability (Bloom et al., 1998), and the present study confirm some, but not all, of the previously reported differences between AMDV-G and AMDV-Utah in this region (Table 2). In addition a novel change in AMDV-Utah (T116L) is reported here.

Amino acid 420 have been proposed to increase viral fitness by prevention of caspase cleavage (Cheng et al., 2010), however in agreement with other studies (Bloom et al., 1988; Oie et al., 1996; Sang et al., 2012) that particular difference between AMDV-G and AMDV-Utah was not observed here either. VP2 amino acid 428–446 functions as a small part of the capsid which has also been suggested to be important for immunopathogenesis by defining AMDV host range (McKenna et al., 1999). The present study confirms a previously reported difference, N343H, in this area, but whether or not this change results in increased pathogenicity is currently unknown (Bloom et al., 1988; Oie et al., 1996; Sang et al., 2012).

In addition to the above-mentioned differences between AMDV-G and AMDV-Utah, additional differences were observed in the VP2 amino acid sequence between the AMDV-G reference and the AMDV-G strain from the laboratory (e.g. K90Q, H395N and D574N). It is currently unknown if there is a fitness effect associated to these changes (e.g. adaption to tissue-culture conditions).

The overall lower conservation of AMDV NS-genes compared to other parvoviruses is supported in the present study by the higher degree of variability in the left ORF compared to in the right ORF, which was even more striking on amino acid level (Fig. 5, panel A).

3.6. Regulatory elements

Previous studies have identified eight TATA-boxes in the AMDV genome; two confirmed functional at nucleotide 154 (TATAA) and 1729 (TATTAA), and six additional boxes at nucleotide 665, 818, 2546, 4136, 4394, and 4468 (AATAAA) with unknown function (Bloom et al., 1988). In the AMDV-Utah sequence, a previously reported difference to AMDV-G in TATA-box 818 (T820C) (Bloom et al., 1988; Gottschalck et al., 1994) is confirmed, and a previously not described change in the 665 box (A669G) is reported. But since the function of these TATA boxes is unknown, the importance of these differences remains to be investigated. The 4468 box was not included in the sequences generated in the present study.

A P3 promotor that initiates transcription of all mRNA have previously identified at nucleotide 151–160 (p3, GTATATAAGC) (Bloom et al., 1988), in addition to a more uncertain promotor P36 around nucleotide 1744 (Bloom et al., 1988; Qiu et al., 2006). In the present study these promotors were fully conserved. However at the suggested transcription initiation site at nucleotide 179 a change (A179G) is reported in the present AMDV-G strain (Table 2) compared to the reference AMDV-G genome.

Internal polyadenylation sites (pA)p's at nucleotide 2561 and 4391 have been suggested to play a major role in AMDV replication (Huang et al., 2012). Studies in the parvovirus Minute Virus of Mice (MVM) shows that the above mentioned NS1 GKRN-region contains a NS1 recognition site at amino acids 337–344 (ACCAACCA), which together with an upstream nicking site initiates viral replication (Christensen et al., 1997). All of these sites were conserved in the two viruses investigated here, and therefore should not have any effect on the increased virulence of AMDV-Utah.

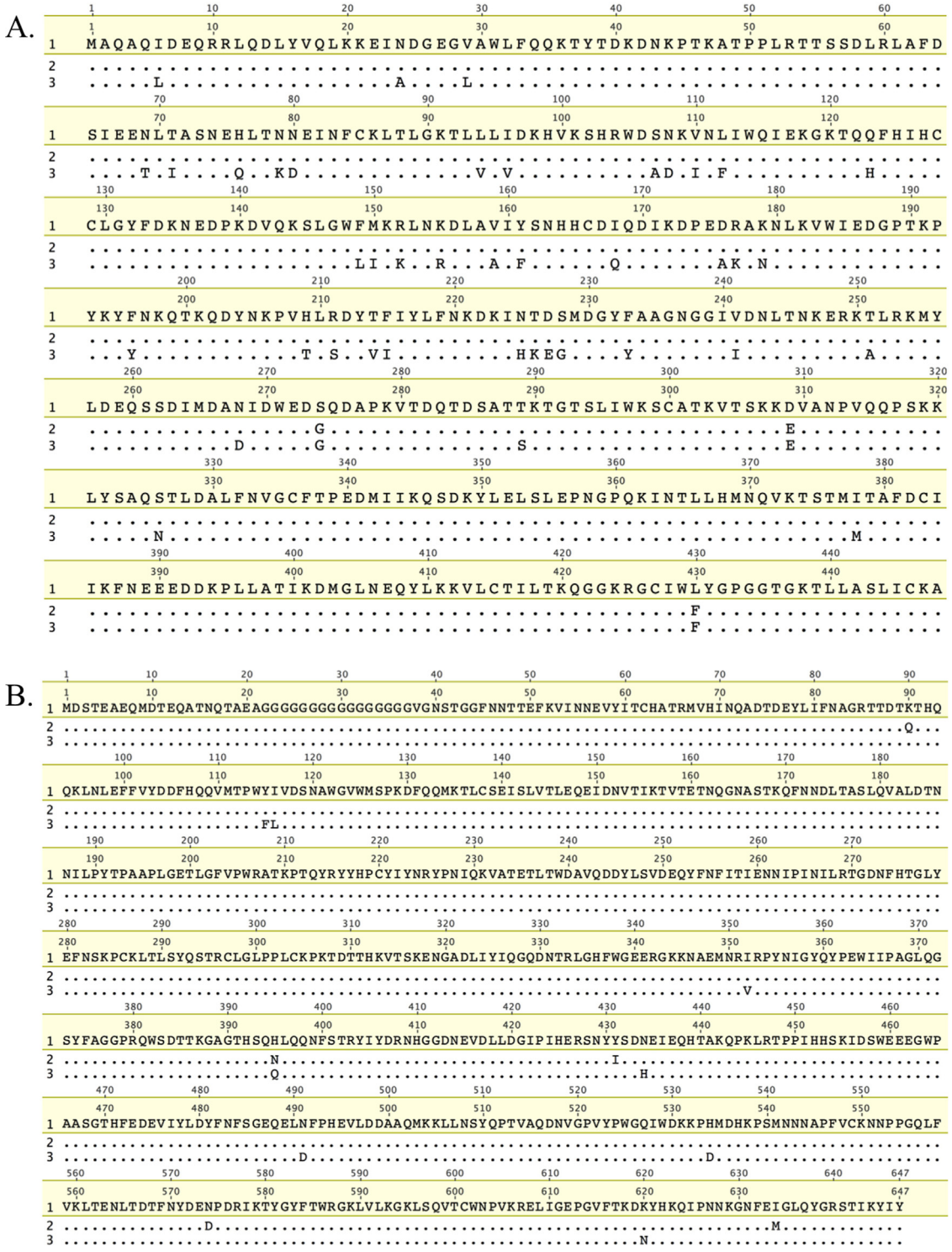

Fig. 5. Protein alignments for each of the two major AMDV genes. Translation and alignments at protein level for each of the NS1 (Panel A) and VP2 (Panel B) genes for the AMDV-G reference NC_001662 (1), one representative AMDV-G strain (2), and the AMDV-Utah (3) strain sequenced in this study.

4. Discussion

This paper describes a fast and robust protocol for next generation sequencing of the near full length AMDV genome and the subsequent data analysis. The protocol was verified by gel electrophoresis, complimentary Sanger sequencing, and by sequence analysis. The prototypic non-virulent cell-culture adapted strain AMDV-G was used as a model virus, and for comparison and to investigate genetic virulence markers the highly virulent AMDV-Utah strain was also sequenced. Due to the presence of secondary structures and palindromic motifs at the 3′ and 5′ ends, approximately 91% of the viral genome was amplified; including all known coding regions (Alexandersen et al., 1988; Bloom et al., 1988).

This study confirm some of the previously reported nucleotide and amino acid differences between AMDV-G and AMDV-Utah (Table 2) and no major deviations in the suggested genomic regulatory regions were observed (Bloom et al., 1988). Therefore, one can speculate that the increased virulence of AMDV-Utah compared

to AMDV-G is not due to differences in gene regulation but rather on protein level.

However, some nucleotide and amino acid differences was observed between the previously published Sanger generated AMDV-G and AMDV-Utah genomes and the genomes sequenced in this study, which could be a result of further cell-culture adaptation of both strains or due to the use of different sequencing technologies. One specific change of interest is the A179 G seen in our AMDV-G strain (Table 2), as it might influence the translation initiation codon at nt 179–181. Another change that might be of importance is the F430L. It resides in close proximity to the conserved ATP-binding pocket and GKRN region between amino acids 435–440, which has been suggested to be essential for the NS1 protein and viral DNA replication due to its ATP- and GTP binding sites and its ATPase function (Gottschalck et al., 1994).

There is some disagreement in the literature regarding the frame and sequence of the first three amino acids of the AMDV-G VP1 gene, in that one study suggests the start to be MSK in frame 2 (Huang et al., 2012) (accession number JN040434.1), while another study suggests HHN in frame 3 (Schuierer et al., 1997) (accession number X97629.1). The first option would be more similar to e.g. human parvovirus B19, in which the VP1-unique region (starting with an MSK) is encoded in another frame, while the remaining protein is identical to that of VP2 (as depicted in Fig. 4). In the second option, starting with amino acids HHN, the whole VP1 protein is encoded in the same frame as that of VP2, but with an additional 55 amino acids in the N-terminal. However, in this sequence there is no start codon (Fig. 4), and therefore it is sensible to assume that option one (MSK) is more correct. In human parvovirus B19 this unique VP1 N-terminal has been identified as key for viral entry (Leisi et al., 2013), and studies in AMDV suggest its importance as it provides phospholipase A2 enzyme activity which modifies the endosome membrane thereby mediating capsid release (Fenner's Veterinary Virology, 2011). It has further been suggested that AMDV's ability to grow in cell culture is regulated by the N-terminal of the VP2 gene (Bloom et al., 1998), and one can therefore speculate that the VP1 unique terminal could be linked to *in vivo* infectivity.

Both the VP and the NS proteins have been suggested to be involved in determining the viral host range and influence pathogenicity (Fields et al., 2007). For example, it has been shown that knockout of the NS1-gene resulted in failure to produce replicative form AMDV DNA (Huang et al., 2014). In the present study, the majority of nucleotide and amino-acid differences between AMDV-G and AMDV-Utah were in the left ORF. This is in agreement with previous studies reporting that the non-structural (NS) proteins of AMDV-G and AMDV-Utah have different molecular weights (Alexandersen et al., 1986), and are less conserved than in other parvoviruses (Gottschalck et al., 1994). Interestingly, it is known that the AMDV right ORF is more conserved despite containing virulence factors important for the viral entry (Gottschalck et al., 1991; Oie et al., 1996). These findings indicate that the virulence of AMDV-Utah may not be primarily due to increased infectivity since this function depends primarily on the VP genes. Instead the difference in virulence could e.g. be linked to the NS proteins that are involved with virion assembly, release, unpacking, or the ability to avoid host cell responses.

The protocols developed in this study enable viral DNA to be extracted and amplified from primary sample material and by that avoiding the use of labour intensive cloning to amplify the viral DNA prior to sequencing. The PCR-amplification step is also useful as the concentration of viral DNA in viraemic animals is not sufficient to directly act as template for next generation sequencing. The viral strains used to establish this protocol have very different phenotypes and despite the expected genetic difference both strains amplified well. This indicates a high degree of conservation in the primer-annealing region, which is further supported by on-going work successfully amplifying AMDV field strains (data not shown). There are however potential biases when using PCR amplified DNA as input for sequencing, e.g. for investigating quasi-species. But this is of less importance if the resulting data will be used for comparing sample consensus sequences, as in the present study. The ion-semiconductor technology is known to have difficulties to accurately PCR amplify and read homopolymers (Quail et al., 2012), and especially G'- and C'-rich regions as between position 2470 and 2520 in the AMDV-G reference genome. Therefore the robustness of the protocols was demonstrated by processing the AMDV-G strain in triplicates and by sequencing the homopolymeric region in each sample using the complimentary Sanger sequencing method. Since the NGS sequences from each sample were identical, and so were the Sanger generated fragments (the latter with the exception of one erroneous base), it was concluded that the dip in coverage was caused during the sequencing, and not by the PCR amplification. Thus, it would be beneficial to disregard this region in alignments when the Ion Torrent technology is used, and for the development of molecular tools.

In conclusion, this is to the authors' knowledge the first study to describe the entire coding sequence of the AMDV genome using next generation sequencing. The study provides a robust and fast method for generating whole genome sequences of AMDV from various DNA sources and will create value by allowing for phylogenetic analysis with higher resolution and by facilitating development of new diagnostic tools.

Acknowledgements

Kopenhagen Diagnostics, Kopenhagen Fur, and Professor emeritus Bent Aasted, UCPH, are gratefully thanked for supplying the viral strains AMDV-G and AMDV-Utah, respectively.

References

Aasted, B., 1980. Purification and characterization of Aleutian disease virus. Acta Pathol. Microbiol. Scand. B 80, 323–328.

Alexandersen, S., Uttenthal-Jensen, A., Aasted, B., 1986. Demonstration of non-degraded aleutian disease virus (ADV) proteins in lung tissue from experimentally infected mink kits. Arch. Virol. 87, 127–133, http://dx.doi.org/10.1007/BF01310549.

Alexandersen, S., Bloom, M.E., Perryman, S., 1988. Detailed transcription map of Aleutian mink disease parvovirus. J. Virol. 62, 3684–3694.

Bloom, M.E., Race, R.E., Wolfinbarger, J.B., 1980. Characterization of Aleutian disease virus as a parvovirus. J. Virol. 35, 836–843.

Bloom, M.E., Alexandersen, S., Perryman, S., Lechner, D., Wolfinbarger, J.B., 1988. Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus (ADV): sequence comparisons between a nonpathogenic and a pathogenic strain of ADV. J. Virol. 62, 2903–2915.

Bloom, M.E., Alexandersen, S., Garon, C.F., Mori, S., Wei, W., Perryman, S., Wolfinbarger, J.B., 1990. Nucleotide sequence of the 5′-terminal palindrome of Aleutian mink disease parvovirus and construction of an infectious molecular clone. J. Virol. 64, 3551–3556.

Bloom, M.E., Fox, J.M., Berry, B.D., Oie, K.L., Wolfinbarger, J.B., 1998. Construction of pathogenic molecular clones of Aleutian mink disease parvovirus that replicate both in vivo and in vitro. Virology 251, 288–296, http://dx.doi.org/10.1006/viro.1998.9426.

Broll, S., Alexandersen, S., 1996. Investigation of the pathogenesis of transplacental transmission of Aleutian mink disease parvovirus in experimentally infected mink. J. Virol. 70, 1455–1466.

Cheng, F., Chen, A.Y., Best, S.M., Bloom, M.E., Pintel, D., Qiu, J., 2010. The capsid proteins of Aleutian mink disease virus activate caspases and are specifically cleaved during infection. J. Virol. 84, 2687–2696, http://dx.doi.org/10.1128/JVI.01917-09.

Christensen, J., Cotmore, S.F., Tattersall, P., 1997. Parvovirus initiation factor PIF: a novel human DNA-binding factor which coordinately recognizes two ACGT motifs. J. Virol. 71, 5733–5741.

Christensen, L.S., Gram-Hansen, L., Chriél, M., Jensen, T.H., 2011. Diversity and stability of Aleutian mink disease virus during bottleneck transitions resulting from eradication in domestic mink in Denmark. Vet. Microbiol. 149, 64–71, http://dx.doi.org/10.1016/j.vetmic.2010.10.016.

Danecek, P., Auton, A., Abecasis, G., Albers, C.A., Banks, E., DePristo, M.A., Handsaker, R.E., Lunter, G., Marth, G.T., Sherry, S.T., McVean, G., Durbin, R., 2011. The variant call format and VCFtools. Bioinformatics 27, 2156–2158, http://dx.doi.org/10.1093/bioinformatics/btr330.

Danish Executive Order 1447 of 15/12/2009, 2009, Danish Executive Order 1447 of 15/12/2009.
Decaro, Nicola, Buonavoglia, C., Ryser-Degiorgis, M.-P., Gortázar, C., 2012. 12. Parvovirus infections. In: Gavier-Widén, D., Duff, J.P., Meredith, A. (Eds.), Infectious Diseases of Wild Mammals and Birds in Europe. Wiley-Blackwell Oxford, UK, pp. 181–285.
Escobar-Gutiérrez, A., Vazquez-Pichardo, M., Cruz-Rivera, M., Rivera-Osorio, P., Carpio-Pedroza, J.C., Ruíz-Pacheco, J.A., Ruiz-Tovar, K., Vaughan, G., 2012. Identification of hepatitis C virus transmission using a next-generation sequencing approach. J. Clin. Microbiol. 50, 1461–1463, http://dx.doi.org/10.1128/JCM.00005-12.
Fenner's Veterinary Virology, 2011. Fenner's Veterinary Virology. Elsevier, http://dx.doi.org/10.1016/B978-0-12-375158-4.00012-2.
Fields, B.N., Knipe, D.M., Howley, P.M., 2007. Fields Virology, 5th ed. Fields Virol.
Gottschalck, E., Alexandersen, S., Cohn, A., Poulsen, L.A., 1991. Nucleotide sequence analysis of Aleutian mink disease parvovirus shows that multiple virus types are present in infected mink. J. Virol. 65, 4378–4386.
Gottschalck, E., Alexandersen, S., Storgaard, T., Bloom, M.E., Aasted, B., 1994. Sequence comparison of the non-structural genes of four different types of Aleutian mink disease parvovirus indicates an unusual degree of variability. Arch. Virol. 138, 213–231, http://dx.doi.org/10.1007/BF01379127.
Henn, M.R., Boutwell, C.L., Charlebois, P., Lennon, N.J., Power, K.A., Macalalad, A.R., Berlin, A.M., Malboeuf, C.M., Ryan, E.M., Gnerre, S., Zody, M.C., Erlich, R.L., Green, L.M., Berical, A., Wang, Y., Casali, M., Streeck, H., Bloom, A.K., Dudek, T., Tully, D., Newman, R., Axten, K.L., Gladden, A.D., Battis, L., Kemper, M., Zeng, Q., Shea, T.P., Gujja, S., Zedlack, C., Gasser, O., Brander, C., Hess, C., Günthard, H.F., Brumme, Z.L., Brumme, C.J., Bazner, S., Rychert, J., Tinsley, J.P., Mayer, K.H., Rosenberg, E., Pereyra, F., Levin, J.Z., Young, S.K., Jessen, H., Altfeld, M., Birren, B.W., Walker, B.D., Allen, T.M., 2012. Whole genome deep sequencing of HIV-1 reveals the impact of early minor variants upon immune recognition during acute infection. PLoS Pathog. 8, e1002529, http://dx.doi.org/10.1371/journal.ppat.1002529.
Huang, Q., Deng, X., Best, S.M., Bloom, M.E., Li, Y., Qiu, J., 2012. Internal polyadenylation of parvoviral precursor mRNA limits progeny virus production. Virology 426, 167–177, http://dx.doi.org/10.1016/j.virol.2012.01.031.
Huang, Q., Luo, Y., Cheng, F., Best, S.M., Bloom, M.E., Qiu, J., 2014. Molecular characterization of the small nonstructural proteins of parvovirus Aleutian mink disease virus (AMDV) during infection. Virology 452–453, 23–31, http://dx.doi.org/10.1016/j.virol.2014.01.005.
Jakhesara, S.J., Bhatt, V.D., Patel, N.V., Prajapati, K.S., Joshi, C.G., 2014. Isolation and characterization of H9N2 influenza virus isolates from poultry respiratory disease outbreak. Springerplus 3, 196, http://dx.doi.org/10.1186/2193-1801-3-196.
Katoh, K., Standley, D.M., 2013. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. 30, 772–780, http://dx.doi.org/10.1093/molbev/mst010.
Kearse, M., Moir, R., Wilson, A., Stones-Havas, S., Cheung, M., Sturrock, S., Buxton, S., Cooper, A., Markowitz, S., Duran, C., Thierer, T., Ashton, B., Meintjes, P., Drummond, A., 2012. Geneious basic: an integrated and extendable desktop software platform for the organization and analysis of sequence data. Bioinformatics 28, 1647–1649, http://dx.doi.org/10.1093/bioinformatics/bts199.
Knuuttila, A., Aaltonen, K., Virtala, A.-M.K., Henttonen, H., Isomursu, M., Leimann, A., Maran, T., Saarma, U., Timonen, P., Vapalahti, O., Sironen, T., 2015. Aleutian mink disease virus in free-ranging mustelids in Finland—a cross-sectional epidemiologic and phylogenetic study. J. Gen. Virol., http://dx.doi.org/10.1099/vir.0.000081.
Kvisgaard, L.K., Hjulsager, C.K., Fahnoe, U., Breum, S.O., Ait-Ali, T., Larsen, L.E., 2013. A fast and robust method for full genome sequencing of porcine reproductive and respiratory syndrome virus (PRRSV) type 1 and type 2. J. Virol. Methods 193, 697–705, http://dx.doi.org/10.1016/j.jviromet.2013.07.019.
Leimann, A., Knuuttila, A., Maran, T., Vapalahti, O., Saarma, U., 2015. Molecular epidemiology of Aleutian mink disease virus (AMDV) in Estonia, and a global phylogeny of AMDV. Virus Res. 199, 55–61, http://dx.doi.org/10.1016/j.virusres.2015.01.011.
Leisi, R., Ruprecht, N., Kempf, C., Ros, C., 2013. Parvovirus B19 uptake is a highly selective process controlled by VP1u, a novel determinant of viral tropism. J. Virol. 87, 13161–13167, http://dx.doi.org/10.1128/JVI.02548-13.
Li, H., 2013. Aligning sequence reads, clone sequences and assembly contigs with BWA-MEM.
Martin, M., 2011. Cutadapt removes adapter sequences from high-throughput sequencing reads. EMBnet J. 17, 10–12, http://dx.doi.org/10.14806/ej.17.1.200.
McKenna, R., Olson, N.H., Chipman, P.R., Baker, T.S., Booth, T.F., Christensen, J., Aasted, B., Fox, J.M., Bloom, M.E., Wolfinbarger, J.B., Agbandje-McKenna, M., 1999. Three-dimensional structure of Aleutian mink disease parvovirus: implications for disease pathogenicity. J. Virol. 73, 6882–6891.
Oie, K.L., Durrant, G., Wolfinbarger, J.B., Martin, D., Costello, F., Perryman, S., Hogan, D., Hadlow, W.J., Bloom, M.E., 1996. The relationship between capsid protein (VP2) sequence and pathogenicity of Aleutian mink disease parvovirus (ADV): a possible role for raccoons in the transmission of ADV infections. J. Virol. 70, 852–861.
Qiu, J., Cheng, F., Burger, L.R., Pintel, D., 2006. The transcription profile of Aleutian mink disease virus in CRFK cells is generated by alternative processing of pre-mRNAs produced from a single promoter. J. Virol. 80, 654–662, http://dx.doi.org/10.1128/JVI.80.2.654-662.2006.
Quail, M.A., Smith, M., Coupland, P., Otto, T.D., Harris, S.R., Connor, T.R., Bertoni, A., Swerdlow, H.P., Gu, Y., 2012. A tale of three next generation sequencing platforms: comparison of Ion Torrent, pacific biosciences and Illumina MiSeq sequencers. BMC Genomics 13, 341, http://dx.doi.org/10.1186/1471-2164-13-341.
Sang, Y., Ma, J., Hou, Z., Zhang, Y., 2012. Phylogenetic analysis of the VP2 gene of Aleutian mink disease parvoviruses isolated from 2009 to 2011 in China. Virus Genes 45, 31–37, http://dx.doi.org/10.1007/s11262-012-0734-9.
Schmieder, R., Edwards, R., 2011. Quality control and preprocessing of metagenomic datasets. Bioinformatics (Oxford, England) 27, 863–864, http://dx.doi.org/10.1093/bioinformatics/btr026.
Schuierer, S., Bloom, M.E., Kaaden, O.R., Truyen, U., Diseases, E., Disease, I., 1997. Sequence analysis of the lymphotropic Aleutian disease parvovirus ADV-SL3. Brief Report, 157–166.
Shackelton, L.A., Hoelzer, K., Parrish, C.R., Holmes, E.C., 2007. Comparative analysis reveals frequent recombination in the parvoviruses. J. Gen. Virol. 88, 3294–3301, http://dx.doi.org/10.1099/vir.0.83255-0.
Ye, J., Coulouris, G., Zaretskaya, I., Cutcutache, I., Rozen, S., Madden, T.L., 2012. Primer-BLAST: a tool to design target-specific primers for polymerase chain reaction. BMC Bioinf. 13, 134, http://dx.doi.org/10.1186/1471-2105-13-134.